

Sponsor & Client: Phillip Chan (pkc@fit.edu)
26 September 2018

1. **Project Tile & Members**

1.1. **Title**

DengAI

1.2. **Members**

Micah Oltmann (moltmann2015@my.fit.edu)

Hy Diep (hdiep2015@my.fit.edu)

2. **Faculty Sponsor**

Phillip Chan (pkc@fit.edu)

3. **Client**

Phillip Chan (pkc@fit.edu)

4. **Client Meetings**

9/7, 9/14, 9/28

5. **Sponsor Meetings**

Same as Client Meetings.

6. **Current Milestone Progress**

Task	Completion %	Micah	Hy	To do
1. Investigate tools	100%	65%	35%	N/A
2. Hello World demos	95%	95%	0%	N/A
3. Requirement Document	100%	10%	90%	N/A
4. Design Document	100%	100%	0%	N/A
5. Test Plan	100%	10%	90%	user studies

7. **Milestone 1 Task Discussions**

7.1. **Task 1**

Multiple tools are able to do forecasting and modeling. However, due to a discovery of (what we think) is a great intro to forecasting book (Hyndman & Athanasopoulos, *Forecasting: principles and practice*, 2018) we've decided to stick with R since the author also has created a package with examples and example data. We did look at MATLAB: Econometrics Toolbox, and though it can do the same things, it's missing sample data sets and is expensive.

7.2. **Task 2**

A simple demo showing that data from a CSV file can be read and outputted by R, furthermore a python wrapper was created to run the R program. No issues were had when creating the demo.

7.3. **Task 3**

Through multiple iterations on what we thought was a simple requirements documents unfolded into many different specifications. From various modeling techniques and data manipulation to various tests to improve the models fitness and therefore the underlying forecasting accuracy. Lots of researching and reading was needed to be able understand even the basics of what modeling and forecasting is. Furthermore, the help from Dr. Chan's knowledge of forecasting helped a great deal with pointing us towards the correct direction.

7.4. **Task 4**

Due to little little expertise in creating design documents, our first draft wasn't our best job. However, Dr. Chan was a great help in both leading us and showing us some ideas for creating this document. In the end, through multiple iterations, and the great help of Dr. Chan, we've produced a passable and good design document.

7.5. **Task 5**

Once the requirements document was finished, the tests for each of the modules were simple to create as the testing involves comparing sample data from various sources and the output thereof in our system. The harder part was developing the internal validation system which goes back to a Task 3 -- a lot of research and reading into forecasting was required.

8. **Member Contributions**

8.1. **Both**

Worked on selecting tools for this program. We've decided to choose R, Python for our forecasting and PyGUI as a frontend solution.

8.2. **Hy Diep**

Worked on creating the requirement documents and test documents with input from Dr. Chan. Multiple revisions have been made, each time including more

requirements such as various internal validation schemes, input specifications, and the required tests thereof.

8.3. **Micah Oltmann**

Worked on the Design Document with input from Dr. Chan and created a system diagram showing what modules are in the system, how the users interact with this system and how the data is used and transferred between the users and internal modules. Furthermore, a sketch of the GUI system was created.

9. **Next Milestone**

Task	Micah	Hy
1. Calculating Pearson Correlation Coefficients for each input variable	Inputting the data and making sure the data is stored in the correct formats.	Run pearson correlation calculations on the inputted data.
2. Simple regression modeling with top K most correlated predictors and forecasts thereof	Creating the forecasts	Creating the models
3. Internal validation system including multiple diagnostic plots and splitting of data	Creating plots of errors, forecasted values, actual values and plots of residuals vs predictors, and calculating the MAE	Calculating residuals, ACF Plots and running the Breusch-Godfrey test
4. Starting the Automated Workflow System	Working on building a simple GUI to select between the different options.	Help streamline functions to be able to be used for the workflow system.

10. Milestone 2 Tasks

10.1. Task 1

For each of the predictors input to the system, pearson correlation coefficients are be calculated between the predictor and dengue fever cases. The top K correlated predictors are used in the next task.

10.2. Task 2

With each of the top K correlated predictors, simple linear regression models are generated with the dengue fever cases. Forecasts are then generated with each of those models.

10.3. Task 3

An internal validation system is created to determine the fitness of the model, and the accuracy of the model. This system splits the input data into training and test data. This is used to help evaluate this milestone's forecasts and models as well as the following milestones. These methods include ACF Plots, the Breusch-Godfrey test, residuals plotting and MAE calculations.

10.4. Task 4

Although all the bits and pieces of this system are spread throughout the other milestones, we are starting to piece together some of the fundamental tasks we're doing in this milestone to create a GUI which will expand as more features are added.

11. Sponsor Feedback

11.1. Task 1

Investigate more tools (technical and collaboration) and determine the pros and cons thereof.

11.2. Task 2

N/A

11.3. Task 3

Remove duplicate mentions of functions and combine program functions with requirements. Rename some subsections (Automated Workflow -> User Center workflow system, Efficiency -> Fitness, Accuracy -> Forecast's Accuracy). Addition of more requirements for nonlinear systems, differentiated predictors, and time-lagged predictors.

11.4. Task 4

Reconfigure the UML diagram to better represent the user interaction with the automated workflow and the ability to specify forecasting method, internal validation, and plotting.

11.5. Task 5: Reorganize test cases to follow structure outlined in the requirements document. Relabel sections to better fit the requirements document.

12. **Sponsor Feedback**

Micah Oltmann	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Hy Diep	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10

Sponsor Signature: _____ Date: _____